

OICE

## RAW SEQUENCE LISTING

DATE: 08/15/2001

PATENT APPLICATION: US/09/667,569A

TIME: 14:45:51

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\08152001\I667569A.raw

ENTERED

3 <110> APPLICANT: Yocum, R. Rogers  
4 Patterson, Thomas A.  
5 Hermann, Theron  
6 Pero, Janice G.  
8 <120> TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR PRODUCTION OF  
9 PANTO-COMPOUNDS  
11 <130> FILE REFERENCE: BGI-141CP  
13 <140> CURRENT APPLICATION NUMBER: US/09/667,569A  
14 <141> CURRENT FILING DATE: 2000-09-21  
16 <150> PRIOR APPLICATION NUMBER: USSN 09/400,494  
17 <151> PRIOR FILING DATE: 1999-09-21  
19 <150> PRIOR APPLICATION NUMBER: USSN 60/210,072  
20 <151> PRIOR FILING DATE: 2000-06-07  
22 <150> PRIOR APPLICATION NUMBER: USSN 60/221,938  
23 <151> PRIOR FILING DATE: 2000-07-28  
25 <150> PRIOR APPLICATION NUMBER: USSN 60/227,860  
26 <151> PRIOR FILING DATE: 2000-08-24  
28 <160> NUMBER OF SEQ ID NOS: 94  
30 <170> SOFTWARE: PatentIn Ver. 2.0  
32 <210> SEQ ID NO: 1  
33 <211> LENGTH: 311  
34 <212> TYPE: PRT  
35 <213> ORGANISM: Haemophilus influenzae  
37 <400> SEQUENCE: 1  
38 Met Glu Phe Ser Thr Gln Gln Thr Pro Phe Leu Ser Phe Asn Arg Glu  
39 1 5 10 15  
41 Gln Trp Ala Glu Leu Arg Lys Ser Val Pro Leu Lys Leu Thr Glu Gln  
42 20 25 30  
44 Asp Leu Lys Pro Leu Leu Gly Phe Asn Glu Asp Leu Ser Leu Asp Glu  
45 35 40 45  
47 Val Ser Thr Ile Tyr Leu Pro Leu Thr Arg Leu Ile Asn Tyr Tyr Ile  
48 50 55 60  
50 Asp Glu Asn Leu His Arg Gln Thr Val Leu His Arg Phe Leu Gly Arg  
51 65 70 75 80  
53 Asn Asn Ala Lys Thr Pro Tyr Ile Ile Ser Ile Ala Gly Ser Val Ala  
54 85 90 95  
56 Val Gly Lys Ser Thr Ser Ala Arg Ile Leu Gln Ser Leu Leu Ser His  
57 100 105 110  
59 Trp Pro Thr Glu Arg Lys Val Asp Leu Ile Thr Thr Asp Gly Phe Leu  
60 115 120 125  
62 Tyr Pro Leu Asn Lys Leu Lys Gln Asp Asn Leu Leu Gln Lys Lys Gly  
63 130 135 140  
65 Phe Pro Val Ser Tyr Asp Thr Pro Lys Leu Ile Arg Phe Leu Ala Asp  
66 145 150 155 160  
68 Val Lys Ser Gly Lys Ser Asn Val Thr Ala Pro Ile Tyr Ser His Leu  
69 165 170 175  
71 Thr Tyr Asp Ile Ile Pro Asp Lys Phe Asp Val Val Asp Lys Pro Asp

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/667,569A

DATE: 08/15/2001

TIME: 14:45:51

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\08152001\I667569A.raw

```

72          180          185          190
74 Ile Leu Ile Leu Glu Gly Leu Asn Val Leu Gln Thr Gly Asn Asn Lys
75          195          200          205
77 Thr Asp Gln Thr Phe Val Ser Asp Phe Val Asp Phe Ser Ile Tyr Val
78          210          215          220
80 Asp Ala Glu Glu Lys Leu Leu Lys Glu Trp Tyr Ile Lys Arg Phe Leu
81 225          230          235          240
83 Lys Phe Arg Glu Ser Ala Phe Asn Asp Pro Asn Ser Tyr Phe Lys His
84          245          250          255
86 Tyr Ala Ser Leu Ser Lys Glu Glu Ala Ile Ala Thr Ala Ser Lys Ile
87          260          265          270
89 Trp Asp Glu Ile Asn Gly Leu Asn Leu Asn Gln Asn Ile Leu Pro Thr
90          275          280          285
92 Arg Glu Arg Ala Asn Leu Ile Leu Lys Lys Gly His Asn His Gln Val
93          290          295          300
95 Glu Leu Ile Lys Leu Arg Lys
96 305          310
99 <210> SEQ ID NO: 2
100 <211> LENGTH: 316
101 <212> TYPE: PRT
102 <213> ORGANISM: Escherichia coli
104 <400> SEQUENCE: 2
105 Met Ser Ile Lys Glu Gln Thr Leu Met Thr Pro Tyr Leu Gln Phe Asp
106 1          5          10          15
108 Arg Asn Gln Trp Ala Ala Leu Arg Asp Ser Val Pro Met Thr Leu Ser
109          20          25          30
111 Glu Asp Glu Ile Ala Arg Leu Lys Gly Ile Asn Glu Asp Leu Ser Leu
112          35          40          45
114 Glu Glu Val Ala Glu Ile Tyr Leu Pro Leu Ser Arg Leu Leu Asn Phe
115          50          55          60
117 Tyr Ile Ser Ser Asn Leu Arg Arg Gln Ala Val Leu Glu Gln Phe Leu
118 65          70          75          80
120 Gly Thr Asn Gly Gln Arg Ile Pro Tyr Ile Ile Ser Ile Ala Gly Ser
121          85          90          95
123 Val Ala Val Gly Lys Ser Thr Thr Ala Arg Val Leu Gln Ala Leu Leu
124          100          105          110
126 Ser Arg Trp Pro Glu His Arg Arg Val Glu Leu Ile Thr Thr Asp Gly
127          115          120          125
129 Phe Leu His Pro Asn Gln Val Leu Lys Glu Arg Gly Leu Met Lys Lys
130          130          135          140
132 Lys Gly Phe Pro Glu Ser Tyr Asp Met His Arg Leu Val Lys Phe Val
133 145          150          155          160
135 Ser Asp Leu Lys Ser Gly Val Pro Asn Val Thr Ala Pro Val Tyr Ser
136          165          170          175
138 His Leu Ile Tyr Asp Val Ile Pro Asp Gly Asp Lys Thr Val Val Gln
139          180          185          190
141 Pro Asp Ile Leu Ile Leu Glu Gly Leu Asn Val Leu Gln Ser Gly Met
142          195          200          205
144 Asp Tyr Pro His Asp Pro His His Val Phe Val Ser Asp Phe Val Asp

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/667,569A

DATE: 08/15/2001

TIME: 14:45:52

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\08152001\I667569A.raw

```

145      210      215      220
147 Phe Ser Ile Tyr Val Asp Ala Pro Glu Asp Leu Leu Gln Thr Trp Tyr
148 225      230      235      240
150 Ile Asn Arg Phe Leu Lys Phe Arg Glu Gly Ala Phe Thr Asp Pro Asp
151      245      250      255
153 Ser Tyr Phe His Asn Tyr Ala Lys Leu Thr Lys Glu Glu Ala Ile Lys
154      260      265      270
156 Thr Ala Met Thr Leu Trp Lys Glu Ile Asn Trp Leu Asn Leu Lys Gln
157      275      280      285
159 Asn Ile Leu Pro Thr Arg Glu Arg Ala Ser Leu Ile Leu Thr Lys Ser
160      290      295      300
162 Ala Asn His Ala Val Glu Glu Val Arg Leu Arg Lys
163 305      310      315
166 <210> SEQ ID NO: 3
167 <211> LENGTH: 319
168 <212> TYPE: PRT
169 <213> ORGANISM: Bacillus subtilis
171 <400> SEQUENCE: 3
172 Met Lys Asn Lys Glu Leu Asn Leu His Thr Leu Tyr Thr Gln His Asn
173 1 5 10 15
175 Arg Glu Ser Trp Ser Gly Phe Gly Gly His Leu Ser Ile Ala Val Ser
176 20 25 30
178 Glu Glu Glu Ala Lys Ala Val Glu Gly Leu Asn Asp Tyr Leu Ser Val
179 35 40 45
181 Glu Glu Val Glu Thr Ile Tyr Ile Pro Leu Val Arg Leu Leu His Leu
182 50 55 60
184 His Val Lys Ser Ala Ala Glu Arg Asn Lys His Val Asn Val Phe Leu
185 65 70 75 80
187 Lys His Pro His Ser Ala Lys Ile Pro Phe Ile Ile Gly Ile Ala Gly
188 85 90 95
190 Ser Val Ala Val Gly Lys Ser Thr Thr Ala Arg Ile Leu Gln Lys Leu
191 100 105 110
193 Leu Ser Arg Leu Pro Asp Arg Pro Lys Val Ser Leu Ile Thr Thr Asp
194 115 120 125
196 Gly Phe Leu Phe Pro Thr Ala Glu Leu Lys Lys Lys Asn Met Met Ser
197 130 135 140
199 Arg Lys Gly Phe Pro Glu Ser Tyr Asp Val Lys Ala Leu Leu Glu Phe
200 145 150 155 160
202 Leu Asn Asp Leu Lys Ser Gly Lys Asp Ser Val Lys Ala Pro Val Tyr
203 165 170 175
205 Ser His Leu Thr Tyr Asp Arg Glu Glu Gly Val Phe Glu Val Val Glu
206 180 185 190
208 Gln Ala Asp Ile Val Ile Ile Glu Gly Ile Asn Val Leu Gln Ser Pro
209 195 200 205
211 Thr Leu Glu Asp Asp Arg Glu Asn Pro Arg Ile Phe Val Ser Asp Phe
212 210 215 220
214 Phe Asp Phe Ser Ile Tyr Val Asp Ala Glu Glu Ser Arg Ile Phe Thr
215 225 230 235 240
217 Trp Tyr Leu Glu Arg Phe Arg Leu Leu Arg Glu Thr Ala Phe Gln Asn

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/667,569A

DATE: 08/15/2001

TIME: 14:45:52

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\08152001\I667569A.raw

```

218                               245                               250                               255
220 Pro Asp Ser Tyr Phe His Lys Phe Lys Asp Leu Ser Asp Gln Glu Ala
221                               260                               265                               270
223 Asp Glu Met Ala Ala Ser Ile Trp Glu Ser Val Asn Arg Pro Asn Leu
224                               275                               280                               285
226 Tyr Glu Asn Ile Leu Pro Thr Lys Phe Arg Ser Asp Leu Ile Leu Arg
227                               290                               295                               300
229 Lys Gly Asp Gly His Lys Val Glu Glu Val Leu Val Arg Arg Val
230 305                               310                               315
233 <210> SEQ ID NO: 4
234 <211> LENGTH: 312
235 <212> TYPE: PRT
236 <213> ORGANISM: Mycobacterium leprae
238 <400> SEQUENCE: 4
239 Met Pro Arg Leu Ser Glu Pro Ser Pro Tyr Val Glu Phe Asp Arg Lys
240 1                               5                               10                               15
242 Gln Trp Arg Ala Leu Arg Met Ser Thr Pro Leu Ala Leu Thr Glu Glu
243                               20                               25                               30
245 Glu Leu Ile Gly Leu Arg Gly Leu Gly Glu Gln Ile Asp Leu Leu Glu
246                               35                               40                               45
248 Val Glu Glu Val Tyr Leu Pro Leu Ala Arg Leu Ile His Leu Gln Val
249 50                               55                               60
251 Ala Ala Arg Gln Arg Leu Phe Ala Ala Thr Ala Glu Phe Leu Gly Glu
252 65                               70                               75                               80
254 Pro Gln Gln Asn Pro Gly Arg Pro Val Pro Phe Ile Ile Gly Val Ala
255                               85                               90                               95
257 Gly Ser Val Ala Val Gly Lys Ser Thr Thr Ala Arg Val Leu Gln Ala
258                               100                              105                              110
260 Leu Leu Ala Arg Trp Asp His His Thr Arg Val Asp Leu Val Thr Thr
261                               115                              120                              125
263 Asp Gly Phe Leu Tyr Pro Asn Ala Glu Leu Gly Arg Arg Asn Leu Met
264 130                              135                              140
266 His Arg Lys Gly Phe Pro Glu Ser Tyr Asn Arg Arg Ala Leu Met Arg
267 145                              150                              155                              160
269 Phe Val Thr Ser Val Lys Ser Gly Ala Asp Tyr Ala Cys Ala Pro Val
270                               165                              170                              175
272 Tyr Ser His Leu Arg Tyr Asp Thr Ile Pro Gly Ala Lys His Val Val
273                               180                              185                              190
275 Arg His Pro Asp Ile Leu Ile Leu Glu Gly Leu Asn Val Leu Gln Thr
276                               195                              200                              205
278 Gly Pro Thr Leu Met Val Ser Asp Leu Phe Asp Phe Ser Leu Tyr Val
279 210                              215                              220
281 Asp Ala Arg Ile Gln Asp Ile Glu Gln Trp Tyr Val Ser Arg Phe Leu
282 225                              230                              235                              240
284 Ala Met Arg Gly Thr Ala Phe Ala Asp Pro Glu Ser His Phe His His
285                               245                               250                               255
287 Tyr Ser Ala Leu Thr Asp Ser Lys Ala Ile Ile Ala Ala Arg Glu Ile
288                               260                               265                               270
290 Trp Arg Ser Ile Asn Arg Pro Asn Leu Val Glu Asn Ile Leu Pro Thr

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/667,569A

DATE: 08/15/2001

TIME: 14:45:52

Input Set : A:\seqlistcorrected.txt

Output Set: N:\CRF3\08152001\I667569A.raw

```

291          275          280          285
293 Arg Pro Arg Ala Thr Leu Val Leu Arg Lys Asp Ala Asp His Ser Ile
294          290          295          300
296 Asn Arg Leu Arg Leu Arg Lys Leu
297 305          310
300 <210> SEQ ID NO: 5
301 <211> LENGTH: 312
302 <212> TYPE: PRT
303 <213> ORGANISM: Mycobacterium tuberculosis
305 <400> SEQUENCE: 5
306 Met Ser Arg Leu Ser Glu Pro Ser Pro Tyr Val Glu Phe Asp Arg Arg
307 1          5          10          15
309 Gln Trp Arg Ala Leu Arg Met Ser Thr Pro Leu Ala Leu Thr Glu Glu
310          20          25          30
312 Glu Leu Val Gly Leu Arg Gly Leu Gly Glu Gln Ile Asp Leu Leu Glu
313          35          40          45
315 Val Glu Glu Val Tyr Leu Pro Leu Ala Arg Leu Ile His Leu Gln Val
316          50          55          60
318 Ala Ala Arg Gln Arg Leu Phe Ala Ala Thr Ala Glu Phe Leu Gly Glu
319 65          70          75          80
321 Pro Gln Gln Asn Pro Asp Arg Pro Val Pro Phe Ile Ile Gly Val Ala
322          85          90          95
324 Gly Ser Val Ala Val Gly Lys Ser Thr Thr Ala Arg Val Leu Gln Ala
325          100          105          110
327 Leu Leu Ala Arg Trp Asp His His Pro Arg Val Asp Leu Val Thr Thr
328          115          120          125
330 Asp Gly Phe Leu Tyr Pro Asn Ala Glu Leu Gln Arg Arg Asn Leu Met
331          130          135          140
333 His Arg Lys Gly Phe Pro Glu Ser Tyr Asn Arg Arg Ala Leu Met Arg
334 145          150          155          160
336 Phe Val Thr Ser Val Lys Ser Gly Ser Asp Tyr Ala Cys Ala Pro Val
337          165          170          175
339 Tyr Ser His Leu His Tyr Asp Ile Ile Pro Gly Ala Glu Gln Val Val
340          180          185          190
342 Arg His Pro Asp Ile Leu Ile Leu Glu Gly Leu Asn Val Leu Gln Thr
343          195          200          205
345 Gly Pro Thr Leu Met Val Ser Asp Leu Phe Asp Phe Ser Leu Tyr Val
346          210          215          220
348 Asp Ala Arg Ile Glu Asp Ile Glu Gln Trp Tyr Val Ser Arg Phe Leu
349 225          230          235          240
351 Ala Met Arg Thr Thr Ala Phe Ala Asp Pro Glu Ser His Phe His His
352          245          250          255
354 Tyr Ala Ala Phe Ser Asp Ser Gln Ala Val Val Ala Ala Arg Glu Ile
355          260          265          270
357 Trp Arg Thr Ile Asn Arg Pro Asn Leu Val Glu Asn Ile Leu Pro Thr
358          275          280          285
360 Arg Pro Arg Ala Thr Leu Val Leu Arg Lys Asp Ala Asp His Ser Ile
361          290          295          300
363 Asn Arg Leu Arg Leu Arg Lys Leu

```

**Please Note:**  
 Use of n and/or Xaa have been detected in the Sequence Listing. Please review the  
 Sequence Listing to ensure that a corresponding explanation is presented in the <220> to  
 <223> fields of each sequence which presents at least one n or Xaa.

# VERIFICATION SUMMARY

PATENT APPLICATION: US/09/667,569A

DATE: 08/15/2001  
TIME: 14:45:53

Input Set : A:\seqlistcorrected.txt  
Output Set: N:\CRF3\08152001\I667569A.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:2900 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:45  
L:2917 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46  
L:2935 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47  
L:2952 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48  
L:7791 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:88